Submitted to *INFORMS Journal on Computing*
manuscript (Please, provide the manuscript number!)

Optimal Sampling for Simulated Annealing under Noise

Robin C. Ball
Department of Physics, University of Warwick, Coventry, CV47AL, UK, R.C.Ball@warwick.ac.uk
Juergen Branke
Warwick Business School, University of Warwick, Coventry, CV47AL, UK, juergen.branke@wbs.ac.uk
Stephan Meisel
Münster School of Business and Economics, University of Münster, Germany, stephan.meisel@uni-muenster.de

This paper proposes a Simulated Annealing variant for optimization problems in which the solution quality can only be estimated by sampling from a random distribution. The aim is to find the solution with the best expected performance, as for example is typical for problems where solutions are evaluated using a stochastic simulation. Assuming Gaussian noise with known standard deviation, we derive a fully sequential sampling procedure and decision rule. The procedure starts with a single sample of the value of a proposed move to a neighboring solution and then continues to draw more samples until it is able to make a decision to finally accept or reject the move. Under constraints of equilibrium detailed balance at each draw, we find a decoupling between the acceptance criterion and the choice of the rejection criterion. We derive a universally optimal acceptance criterion in the sense of maximizing the acceptance probability per sample, and thus the efficiency of the optimization process. We show that the choice of the move rejection criterion depends on expectations of possible alternative moves and propose a simple and practical (albeit more empirical) solution that still preserves detailed balance. An empirical evaluation shows that the resulting approach is indeed more efficient than several previously proposed Simulated Annealing variants.

Subject classification: Heuristic: Simulated annealing. Simulation Optimization. Simulation: Statistical Analysis. Statistics: Sampling.
Area of review: Simulation

1. Introduction

For many practical real-world optimization problems, a solution's quality can only be estimated from noisy observations, for example if a stochastic simulation is used for evaluation, or if determining the quality involves a physical measurement subject to noise. In this paper, we propose

Optimized Stochastic Annealing (OSA), a variant of Simulated Annealing (SA) for optimization problems under noise with the goal of finding the solution with the best expected performance. Assuming Gaussian noise with known variance, we derive a sampling and decision rule that

1. is fully sequential in the sense that it starts with a single sample and then iteratively draws additional samples until it is confident to make an acceptance or rejection decision for the considered move,

2. is one of very few SA variants developed for noisy problems that obey the fundamental principle of SA, the detailed balance equation, despite the noise. It thus behaves like standard SA under deterministic conditions and will converge to the optimum if the annealing process is sufficiently slow, and

3. is optimal in the sense of maximizing the acceptance probability per sample, thus maximizing the speed of convergence to the thermodynamical equilibrium.

The characteristics of the derived OSA method are discussed and it is empirically compared to two other methods from the literature, the method by Ceperley and Dewing (1999) and SANE proposed by Branke et al. (2008). We show that OSA has better search characteristics and that it outperforms the other methods also on the example of a noisy traveling salesperson problem.

The paper is structured as follows. The basic principles of SA and the effect of noise are explained in Section 2. Section 3 surveys related literature. Section 4 briefly describes the two algorithms we selected as benchmarks. OSA is derived in Section 5 and empirically evaluated in Section 6. The paper concludes with a summary and some suggestions for future work.

2. Simulated Annealing and Noise

2.1. Simulated Annealing

SA is a stochastic metaheuristic inspired by physical annealing processes, where a solid is melted and then carefully cooled down in order to obtain perfect molecular structures, corresponding to a state of minimum energy. It was originally proposed by Kirkpatrick et al. (1983) and has since been successfully used in a large variety of optimization applications. For an extensive empirical analysis of SA on several combinatorial problems, see, e.g., Johnson et al. (1989, 1991). Algorithm 1 describes the procedure. The algorithm maintains a single current solution x_c. In every iteration, a new candidate solution x_n is generated in the neighborhood of the current solution, $N(x_c)$. Then, the algorithm decides whether it accepts the new solution, in which case x_n becomes the new current solution, or not, in which case the new solution is simply discarded. The acceptance probability for a move, $P_A(\Delta)$, depends on the quality difference between current and new solution $\Delta = E(x_n) - E(x_c)$, where $E(x)$ is the cost of solution x (in SA often called energy level) to be minimized, and also T, a parameter called temperature (although we omit this in our notation for

brevity). Over the run, T, and thus the probability to accept worse candidate solutions, is gradually reduced according to a so-called "annealing schedule". The probabilistic acceptance distinguishes SA from a simple local hill-climber and allows it to escape local optima.

Algorithm 1 Simulated Annealing

```
T ← T_0
Generate initial solution x_c
repeat
  repeat
    Generate x_n ∈ N(x_c)          // generate new solution
    Δ ← (E(x_n) − E(x_c))
    with probability P_A(Δ)
      x_c ← x_n                    // accept new solution
  until time to reduce temperature
  T ← r(T)
until termination condition met
```

A key design decision is the annealing schedule, i.e., the starting temperature T_0 and the method to reduce T over the iterations. Often geometric annealing, i.e., the multiplication of the temperature with a factor $0 < \alpha < 1$, or $r(T) = \alpha T$, is used in practice. We refrain here from a discussion about the optimal annealing schedule as our focus is on developing a method that follows a given annealing schedule despite the noise with a minimal number of samples. Instead, the interested reader is referred to van Laarhoven and Aarts (1987), Rees and Ball (1987), Nourani and Andresen (1998), Cohn and Fielding (1999). However, it should be noted that because OSA obeys the detailed balance equation even in noisy environments, just like the standard SA in a deterministic setting, it will reach equilibrium at each temperature level if given sufficient time, and thus converge in probability to the optimal solution if annealing is sufficiently slow.

2.2. Equilibrium and Acceptance Probability

According to statistical mechanics, a system is said to be in thermodynamical equilibrium, if its probability of being in a given state x, $P(x)$, follows the Boltzmann-Gibbs-Distribution:

$$P(x) \propto e^{-\beta E(x)} \tag{1}$$

where $\beta = 1/(k_B T)$ is a factor depending on the temperature, T, and k_B is the Boltzmann constant. A sufficient condition for equilibrium is the so-called detailed balance equation,

$$p_{x_1,x_2} P(x_1) = p_{x_2,x_1} P(x_2) \quad \forall x_1, x_2 \tag{2}$$

with p_{x_1,x_2} denoting the probability that the state changes from solution x_1 to solution x_2. The detailed balance condition then constrains the transition probabilities to obey

$$\frac{p_{x_1,x_2}}{p_{x_2,x_1}} = \frac{P(x_2)}{P(x_1)} = \frac{e^{-\beta E(x_2)}}{e^{-\beta E(x_1)}} = e^{-\beta\Delta}, \tag{3}$$

with $\Delta = E(x_2) - E(x_1)$.

In optimization, the transition probability p_{x_1,x_2} is just the probability to attempt a move from x_1 to x_2, $P_m(x_1 \to x_2)$, times the probability $P_A(\Delta)$ of accepting the move, i.e., $p_{x_1,x_2} = P_m(x_1 \to x_2)P_A(\Delta)$, and thus

$$\frac{P_A(\Delta)P_m(x_1 \to x_2)}{P_A(-\Delta)P_m(x_2 \to x_1)} = e^{-\beta\Delta} \tag{4}$$

with $\beta = 1/T$ (in optimization, the temperature T is only a parameter and can subsume the Boltzmann constant k_B). Assuming that $P_m(x_1 \to x_2) = P_m(x_2 \to x_1)$, which is true for most applications of SA, this simplifies to

$$\frac{P_A(\Delta)}{P_A(-\Delta)} = e^{-\beta\Delta}. \tag{5}$$

Various acceptance probability functions $P_A(\Delta)$ satisfy Equation (5). Metropolis et al. (1953) suggested the so-called *Metropolis* criterion, which is defined as

$$P_A^{Metropolis}(\Delta) = \begin{cases} 1 & : \quad \Delta \leq 0 \\ e^{-\Delta/T} & : \quad \Delta > 0 \end{cases} \tag{6}$$

Another valid acceptance probability function has been proposed by Glauber (1963):

$$P_A^{Glauber}(\Delta) = \frac{1}{1 + e^{\Delta/T}} \tag{7}$$

It has been shown that a higher probability to accept a move leads to a faster convergence of the algorithm to the thermodynamical equilibrium (Hastings 1970, Peskun 1973). Romeo and Sangiovanni-Vincentelli (1991) show that the Metropolis criterion is the best choice for the acceptance criterion in the sense that, for fixed T, the underlying Markov chain has the fastest rate of convergence to the equilibrium. It is thus the most commonly used acceptance criterion in the literature for deterministic problems.

2.3. The Effect of Noise

If evaluations are disturbed by an additive Gaussian noise then the observed quality difference δ is a random variable

$$\delta \sim N(\Delta, \sigma^2) \tag{8}$$

where Δ is the true (underlying) energy difference and σ^2 the variance of the observed energy difference. If δ instead of Δ is used to decide with what probability to accept the move we need to distinguish the *effective* acceptance probability of a move, $P_A(\Delta)$, and the acceptance probability function of the acceptance rule, $A(\delta)$. The former can be computed from the later by

$$P_A(\Delta) = \int_{x=-\infty}^{x=\infty} A(x) f(x-\Delta) dx \tag{9}$$

where $f(z)$ is the pdf of the noise.

Obviously, this seriously impacts SA's ability to make proper acceptance decisions. As has been shown for example in Branke et al. (2008), a decreasing signal-to-noise ratio leads to slower convergence to a worse solution quality level up to the point when search becomes a purely random walk in the limit. To reduce the impact of noise, one possibility is to average over a number of samples of the energy difference. Since this is computationally very expensive, it is important to find ways to restore the algorithm's ability to make correct acceptance decisions as efficiently as possible.

3. Related Work

Kushner (1987) was the first to examine the behavior of SA when the objective function is subject to noise. Based on the theory of large deviations and assuming Gaussian noise, he provides a full asymptotic analysis of the system under suitable conditions. Since then, a number of authors have proposed variants of SA suitable for optimization in noisy environments. This section surveys the previous work, categorized into approaches that increase the number of samples per evaluation (Section 3.1), approaches that memorize all collected data (Section 3.2), and approaches that adjust the acceptance criterion (Section 3.3). We restrict the survey to the single-objective case, for a few papers on multi-objective optimization under noise the interested reader is referred to Alrefaei and Diabat (2009), Gutjahr (2005), Avello et al. (2004), Mattila et al. (2012). Section 4 describes in more detail the two methods we chose as benchmarks.

3.1. Increasing the Number of Samples per Evaluation

3.1.1. Convergence Results One way to guarantee convergence of SA in noisy environments is to reduce the effective noise by averaging each evaluation over multiple samples. Gelfand and Mitter (1989) consider discrete search spaces and use Markov chains to prove that SA converges in probability to the globally optimal solutions if the standard deviation of the noise decreases linearly with the temperature. This means that for constant noise, the number of samples per solution has to increase quadratically with the inverse of the temperature.

For the same setting, Gutjahr and Pflug (1996) specify an annealing schedule for which they prove convergence in probability if the standard deviation decreases at least inversely proportional to the number of iterations. Also, they generalize the convergence proof to any noise distribution that is symmetric and more peaked around zero than the normal distribution.

3.1.2. Sequential Sampling Rather than imposing a pre-defined schedule on how to increase the number of samples during the run, the approaches based on sequential sampling start with a rough quality difference estimate from very few samples, but then iteratively draw additional samples if deemed necessary based on the information gathered so far.

Bulgak and Sanders (1988) keep sampling the quality of the *candidate* solution until the energy difference is statistically significant. As a criterion for statistical significance, they use a confidence interval for $E(x_c)$. They acknowledge that the convergence proof associated with the used annealing schedule does not apply to the stochastic case.

Ahmed and Alkhamis (2002) combine SA with a two-stage ranking and selection procedure. In the first stage, means and variances are estimated from a small number of samples of each solution. The number of samples taken in the second stage is then calculated given an accepted error probability in determining the better solution and an indifference zone. Ahmed and Alkhamis assume Gaussian noise with unknown variance and prove that their procedure converges almost surely to the global optimal solution if the accepted error probability and indifference zone go to zero in the limit. They test their method on two problems but provide no comparisons with other approaches.

Branke et al. (2008) combine the idea of always accepting the better solution as in Fink (1998) with a sequential sampling scheme that takes one sample at a time until the estimated probability for the worse solution to have a better estimated quality is less than the desired acceptance probability for the worse solution. As one of our benchmark algorithms, this approach is described in more detail in Section 4.2.

Prudius and Andradottir (2012) test an SA variant that keeps track of all samples taken so far (see next subsection) and takes additional samples if the candidate solution has been visited before, and it is sufficiently likely that it may be the best solution encountered so far.

3.2. Keeping Track of All Samples

Usually, SA algorithms only maintain information about the current solution. Some authors have proposed to store summary statistics about all the solutions visited during the run.

Fox and Heine (1995) show that SA converges in probability if errors in acceptance decisions only occur during some initial transition time, and argue that this is true also in a noisy environment if each solution is visited and re-evaluated infinitely often, and the average over all evaluations so far is used as performance estimate.

The SA variant proposed in Alrefaei and Andradóttir (1999), instead of returning the last solution visited after termination, returns either the solution visited most often throughout the run or (with better performance) the solution with the best average performance over all samples taken.

Because of this, the algorithm does not require convergence of the underlying Markov chain and the authors suggest to use a constant temperature instead of an annealing process. They prove that, under the condition of decreasing noise, the procedure almost surely converges to the set of optimal solutions. The algorithm is compared to approaches of Gutjahr and Pflug (1996), Gelfand and Mitter (1989), Fox and Heine (1995) and seems to perform best on the test problems considered. The idea to return the solution evaluated most often is also adopted in Ahmed and Alkhamis (2002) and Alkhamis and Ahmed (2004).

Prudius and Andradottir (2012) propose a general framework for simulation optimization algorithms with convergence guarantee that allows for estimating a solution's quality based on all samples taken for this solution, rather than only the samples from the current iteration. They empirically compare three SA variants. For a fixed number of samples per iteration, the variant that estimates performance from all samples performed similar to the variant that estimates performance only based on the current iteration samples. A better performance was achieved by a variant that averaged over all samples but only took additional samples if it was sufficiently likely that the new solution is the best solution encountered so far. A preliminary version of this paper appeared in Prudius and Andradottir (2005).

A drawback of all the approaches in this subsection is that they only work on discrete search spaces and require a lot of memory, having to store quality estimates for *all* visited solutions.

3.3. Modifying the Acceptance Criterion

3.3.1. Approaches Not Obeying the Detailed Balance Equation The SA variant proposed by Alkhamis et al. (1999) uses the lower bound of a confidence interval instead of δ to make acceptance decisions. Thus, the higher the uncertainty about the true quality difference, the lower the probability of accepting the new solution. The authors also provide a convergence proof which is closely related to that of Gutjahr and Pflug (1996) but slightly more general. Experimental evaluation seems to suggest that the SA variant using confidence intervals works slightly better than the approach of Gutjahr and Pflug (1996).

In Alkhamis and Ahmed (2004), the idea of Alkhamis et al. (1999) is combined with the idea from Alrefaei and Andradóttir (1999) of returning the solution visited most often.

Wang and Zhang (2006) assume Gaussian noise with unknown variance and use a hypothesis test to figure out whether the objective function values of the current solution and the candidate solution are significantly different. If this is not the case, the candidate solution is discarded, otherwise, the Metropolis criterion is applied with the estimated mean solution qualities.

Painton and Diwekar (1995) penalize uncertainty by adding a multiple of the estimated standard error to a solution's estimated objective value, similar to Alkhamis et al. (1999). The penalty is

increased with decreasing temperature, thus in the long run the algorithm favors solutions evaluated based on many samples. The number of samples used is modified randomly for each new candidate solution. Since Painton and Diewekar consider a problem with uncertain input parameters rather than noisy objective function values, they can use Latin Hypercube Sampling to reduce the variance of the objective function estimate.

3.3.2. Approaches Obeying the Detailed Balance Equation As explained above, thermodynamical equilibrium and the detailed balance equation (Equations (2) and (5)) are fundamental principles of SA. With noise, the standard algorithms no longer adhere to these principles. An interesting idea is therefore to modify the applied acceptance probability $A(\delta)$ such that the noise-induced effective acceptance probability for a move $P_A(\Delta)$ again obeys the detailed balance equation.

Kennedy and Kuti (1985) derive a parameterized acceptance function that in principle obeys the detailed balance equation. However, the parameter depends on the (a priori usually unknown) range of δ. Furthermore, the wider this range, the lower the resulting acceptance rate and the less efficient the algorithm. The algorithm allows for unknown variance and arbitrary noise distributions.

Ceperley and Dewing (1999) develop two acceptance functions under the assumption of normally distributed noise, one for known and one for unknown variance. They argue that their method is more efficient than Kennedy and Kuti (1985) by providing higher acceptance rates especially for higher levels of noise. In the absence of noise, the given function equals the Metropolis criterion. However, as the variance increases or the temperature is decreased, the resulting acceptance rates drop quickly. More details on the Ceperley and Dewing approach are provided in Section 4.1. Bowler (2001) and Ball et al. (2003a) provide an alternative derivation of the result of Ceperley and Dewing (1999) in the more general context of known but general error distribution.

Two further approaches aim at obeying the detailed balance equation, but deviate slightly. Fink (1998) presents a simple yet powerful threshold acceptance criterion, assuming normally-distributed noise and known variance. The method always accepts the solution with the higher estimated quality, and the appropriate number of samples to make this decision is chosen depending on the temperature. While this method does not meet the thermodynamical equilibrium exactly, it results in acceptance probabilities very close to the Glauber criterion. This approach is also discussed in Ball et al. (2003a,b) and Bowler (2001)

SANE proposed by Branke et al. (2008) starts optimization following the approach by Ceperley and Dewing (1999), but then switches to a method that combines the idea of always accepting the better solution as in Fink (1998) with a sequential sampling scheme. As one of our benchmarks, this method is described in more detail in the following section.

4. Methods Used as Benchmarks

In this section, we describe in more detail the two methods we selected for benchmarking. The first one, proposed by Ceperley and Dewing (1999), has been theoretically derived to follow the detailed balance equation, a property we claim for our method as well. SANE proposed by Branke et al. (2008) is a recent sequential method that has been shown to outperform other promising candidates such as Fink (1998) and Alkhamis et al. (1999).

4.1. Acceptance Criterion by Ceperley and Dewing

Ceperley and Dewing (1999) proposed the following acceptance criterion, which is derived to obey the detailed balance equation for normally distributed noise with known variance σ^2:

$$A(\delta) = \begin{cases} 1 & : \quad \delta \leq -\frac{1}{2}\sigma^2/T \\ e^{-(\delta/T+\frac{1}{2}\sigma^2/T^2)} & : \quad \delta > -\frac{1}{2}\sigma^2/T \end{cases} \tag{10}$$

where δ denotes the observed quality difference.

The main drawback of this acceptance criterion is that it reduces the acceptance probability and thus, for low temperatures, a new solution is only accepted if it is perceived as much better than the current solution. This rarely happens, and thus the algorithm stalls, which will be seen later in the experimental results.

Ceperley and Dewing (1999) also discuss the possibility of reducing the noise by sampling each solution multiple times, and thus allowing for higher acceptance rates at the expense of higher computational effort for each comparison. Below, we therefore also test the above acceptance criterion with $n = 10$ and $n = 100$ samples for each comparison, which effectively reduces the variance to $\sigma^2/10$ and $\sigma^2/100$. The method by Ceperley and Dewing will be denoted as CD01, CD10 and CD100 depending on the number of samples used for each comparison.

4.2. SANE

Branke et al. (2008) proposed a sequential sampling algorithm for Simulated Annealing in Noisy Environments, called SANE. In the beginning, for high temperatures, it follows the method of Ceperley and Dewing (1999). For lower temperatures, it switches to a different mode that always selects the perceived better solution, as the approach by Fink (1998), but adjusts the number of samples sequentially. The sequential sampling is outlined in Algorithm 2. The idea is to start with a single sample, and compute the desired acceptance probability of the new solution according to the Glauber acceptance rule, given the current temperature and assuming that the measured quality difference is the true difference. This is then compared to the Bayesian probability of the perceived better solution (i.e., the one with the better mean) actually being the worse solution (P_{err}). If the desired acceptance probability is lower, an additional sample is taken to reduce the Bayesian probability of accidentally selecting the worse solution. Sample estimators are updated and the

Algorithm 2 Sequential sampling in SANE (Branke et al. 2008)

Draw $n = 1$ sample and estimate Δ by δ
$P_{err}(\delta) = \Phi\left(\frac{-|\delta|\sqrt{n}}{\sigma}\right)$
while $P_{err} > P_A^{Glauber}(|\delta|)$ **do**
  Draw another sample ($n \leftarrow n+1$)
  Update δ and P_{err}
end while
Accept better solution

process is repeated until the Bayesian probability of accidentally selecting the worse solution is lower than the desired acceptance probability. The algorithm then selects the better solution. In an empirical comparison, SANE was shown to outperform Ceperley and Dewing (1999), Fink (1998) and Alkhamis et al. (1999). While this method aims to follow the Glauber acceptance criterion and thus to obey the detailed balance equation, as we show later, it deviates to some extent.

5. Optimized Stochastic Annealing

We wish to accept a move probabilistically based on an underlying performance difference Δ, such that we obey the detailed balance equation (5) without knowing the exact value of Δ. We have access to observations δ_i which serve as independent and identically distributed stochastic estimates of Δ with probability density function

$$f(\delta_i - \Delta), \tag{11}$$

and assume f is Gaussian with zero mean and known standard deviation σ.

As discussed before, a higher probability of accepting a move leads to a faster convergence of SA to the thermodynamical equilibrium (Hastings 1970, Peskun 1973). Thus, if we wish to make our algorithm efficient, and assuming that the computational effort is mostly proportional to the number of samples taken (which is certainly justified if sampling means running a computationally expensive simulation), we should aim at minimizing the expected number of samples between acceptance decisions while still obeying the detailed balance equation. This is equivalent to maximizing the reciprocal, i.e., maximizing the probability of acceptance per sample taken which we call η and which is our measure of efficiency. This can be expressed as

$$\eta = \frac{P_A}{n_s} \tag{12}$$

where P_A is the probability of a move being eventually accepted and n_s is the average number of samples taken to decide to accept or reject it.

In the following sections, we present an efficient sampling and decision rule which maximizes η. For each considered move, in each sample step i, the algorithm takes a sample δ_i, and then makes a three way decision whether first (1) the move can now be accepted, or secondly (2) the move can be rejected (in which case a new move is attempted), or else finally (3) a further sample is taken and we decide again. We impose a constraint of detailed balance of move acceptance at each sample step, detailed in Section 5.1, which has the merit that the optimal acceptance rule (derived in Appendix A) is then simple and universal. The optimal rejection criterion (Section 5.2 and Appendix B) is then computable, and in Section 5.3 we show that it can be approximated by a simple threshold which proves close to optimal in a wide range of scenarios and preserves detailed balance.

5.1. Decision Chain, Detailed Balance Condition and Universal Optimal Acceptance Rule

For each proposed new move, we draw at least one successive sample of the cost change δ_i, $i = 1, 2 \ldots$, and it is convenient to introduce cumulative variables:

$$c_0 = 0; \quad c_i = c_{i-1} + \delta_i \,,\, i > 0. \tag{13}$$

At each sample step, after the n'th observation of cumulative cost difference c_n, we make a three way decision, based in principle on the full history to that step,

$$C_n = \{c_1 \ldots c_n\}, \tag{14}$$

with respective probabilities:

$$\begin{array}{l} A(C_n) \text{ to accept} \\ R(C_n) \text{ to reject} \\ S(C_n) \text{ to continue to the next observation,} \end{array} \tag{15}$$

where $A + R + S = 1$ and they could in principle each depend also on n.

The probability density to observe chain history C_n as far as the n'th observation is given by

$$P_n(C_n|\Delta) = \prod_{i=1}^{n-1} \left(f(c_{i+1} - c_i - \Delta) S(C_i)\right) \, f(c_1 - \Delta). \tag{16}$$

We then apply the acceptance probability $A(C_n)$ leading to acceptance probability at the nth step in terms of the underlying Δ evaluating to

$$P_a(\Delta, n) = \int A(C_n) P_n(C_n|\Delta) dc_1 .. dc_n. \tag{17}$$

The total acceptance probability for the proposed move is finally $P_A(\Delta) = \sum_n P_a(\Delta, n)$ and this sum must obey the condition of detailed balance.

We choose to impose the more strict constraint of detailed balance conditions on the move acceptance $P_a(\Delta, n)$ at each step n, and separately for each of the cases where the combination $c_{n-1} + \frac{1}{2}\beta\sigma^2$ is positive and where it is negative. This constraint can be collected together and written symmetrically as

$$P_a(\Delta, n|\,\mathrm{sgn}(c_{n-1} + \frac{1}{2}\beta\sigma^2))\, e^{\beta\Delta/2} = P_a(-\Delta, n|\,\mathrm{sgn}(c_{n-1} + \frac{1}{2}\beta\sigma^2))\, e^{-\beta\Delta/2}. \tag{18}$$

This clearly delivers detailed balance for the total step and hence total move acceptance $P_A(\Delta)$, and moreover it leads to an important decoupling between the choice of best acceptance rule found in Appendix A, which then turns out to be simple and universal, and the choice of threshold for reject vs continue which becomes a separately optimized choice. If we did not have such detailed balance at each step then there could be trade-offs between imbalances at different steps, sensitive to the choices to reject vs continue, leading to the optimization of acceptance and rejection rules becoming non-separable.

We show in Appenidx A that given the above strict detailed balance condition, at every step the move acceptance is maximised by the acceptance rule

$$A(c_n, c_{n-1}) = \min\left(1,\, e^{-2(c_n+\beta\sigma^2/2)(c_{n-1}+\beta\sigma^2/2)/\sigma^2}\right). \tag{19}$$

An important consequence of this acceptance rule is that the first non-positive value of $c_n + \beta\sigma^2/2$ is always accepted, so no later decision ever inherits a prior negative value and only the one case $\mathrm{sgn}(c_{n-1} + \beta\sigma^2/2) = 1$ ever arises in Equation (18). This globally optimal acceptance rule is universal in that it is independent of the distribution of underlying move changes Δ and of the choice of rejection threshold. It is also unique in that any different rule must entail $A(c_n, C_{n-1}) < 1$ for some $c_n < -\beta\sigma^2/2$, and hence lower acceptance for both signs of $\pm(c_n + \beta\sigma^2/2)$. Note that for $n = 1$ we have $c_0 = 0$ and the acceptance function reduces to the Ceperley and Dewing rule $A(c_1) = \min(1, e^{-\beta(c_1+\beta\sigma^2/2)})$.

Our acceptance rule (19) is globally optimal under the constraint that we impose detailed balance at each decision step n, separately for positive and negative $c_{n-1} + \frac{1}{2}\beta\sigma^2$, with the negative cases then never arising because they all get accepted at the previous step. It is also globally optimal within the more restrictive condition of detailed balance at fixed preceding sample history of a given move. We have also considered removing the constraint of separate balance for separate signs of $c_{n-1} + \frac{1}{2}\beta\sigma^2$ and/or for separate steps. Our acceptance rule is at least a local optimum in this wider space of possibilities, because the allowed linear variations would be to lower an acceptance probability from 1 when $(c_n + \beta\sigma^2/2)(c_{n-1} + \beta\sigma^2/2) < 0$, and the deferred gain cannot exceed the acceptance loss due to the non-interaction of subsequent histories as already noted above. However finding any other competing optima requires consideration of acceptance and rejection strategy coupled together, and we cease to have any prospect of the universal nature of the above acceptance rule.

5.2. Optimal Rejection Strategy

Having not accepted the move at a step with $c_n + \beta\sigma^2/2 > 0$, mathematically we have free choice (so far as detailed balance is concerned) whether to reject the move or to continue and draw another sample of its performance. In practice there must come a point where the expected gain in acceptance of a move from pursuing further sample steps is outweighed by their computational cost, limiting the number of future moves. We will see that this trade-off drives the choice of a move rejection threshold c_n^* at each step n, and that in general this threshold depends on the prior distribution we assume for the underlying Δ of future moves.

Let us start out by presuming knowledge that the simulation achieves η accepted moves per sample which is our measure of efficiency. Whether a decision chain is continued beyond the n'th step should be based on whether the forward expectation exceeds this overall rate of return η. We need to know both the expected gain of acceptance probability $a_{n+1}(c_n)$ from subsequent steps if we continue, and also the expected cost as number of further samples $s_{n+1}(c_n)$ if we continue, and then the rejection threshold c_n^* at the n'th step is set by

$$\eta = \frac{a_{n+1}(c_n^*)}{s_{n+1}(c_n^*)}. \tag{20}$$

Each of a and s can be expressed in terms of the contribution from taking the immediate next step plus the expectation from possible steps beyond that. Thus we have

$$a_{n+1}(c_n) = \mathbb{E}(A(c_{n+1}, c_n)) + \mathbb{E}(S(c_{n+1}, c_n) a_{n+2}(c_{n+1})) \tag{21}$$

$$s_{n+1}(c_n) = 1 + \mathbb{E}\left(S(c_{n+1}, c_n) s_{n+2}(c_{n+1})\right) \tag{22}$$

where $A(c_{n+1}, c_n)$ is the acceptance probability as given by Equation (19) and the corresponding continue probability is given by $S(c_{n+1}, c_n) = 1 - A(c_{n+1}, c_n)$ for $-\beta\sigma^2/2 < c_{n+1} < c_{n+1}^*$ and zero otherwise due to acceptance and rejection respectively. The system gives us a backward recurrence down to $n = 1$ at which we have no choice of prior rejection and fixed value of $c_0 = 0$. At this point the expected return from the unsampled move must give us η, so we have

$$\frac{a_1(0)}{s_1(0)} = \eta \tag{23}$$

as a final closing self-consistency condition on the previously assumed value of η.

We show in Appendix B how all this can be explicitly worked through. Figure 1 shows some results and compares them with those of a simpler scheme discussed in the following section. The expectations in Equations (21,22) include averaging over a belief distribution $P_0(\Delta)$ of the underlying change Δ of an attempted move prior to any samples being drawn. In thermal equilibrium

with detailed balance, we must have $P_0(-\Delta) = e^{-\beta\Delta} P_0(\Delta)$ arising from the thermal bias in the population, leading us to introduce the "dethermalized prior"

$$P_{00}(\Delta) = e^{-\beta\Delta/2} P_0(\Delta), \tag{24}$$

which in thermal equilibrium must be symmetric with respect to $\Delta \leftrightarrow -\Delta$.

5.3. Simpler Rejection Strategy

The rejection strategy found above is truly optimal when the cost of computing it is negligible compared to the cost of sampling, but it requires knowledge of the prior distribution of Δ which is usually not available. Happily there is a much simpler strategy of simply rejecting at

$$c^* = 0,$$

which is both motivated by particular cases and compares reasonably in efficiency η as shown in Figure 1. The comparative performance of the optimal stopping rule was computed numerically following the previous section and Appendix B, taking the dethermalized prior to be Gaussian of mean 0 and standard deviation Σ. As can be seen, the efficiencies of the simple and the optimal approach are very similar over a wide range of parameters. The reduction of performance along each curve is due to thermal energy differences becoming harder to resolve above sample noise at low temperatures (large β). The reduction of performance between curves is a simple equilibrium effect: larger Σ means larger move energy differences Δ, which in equilibrium becomes increasingly biased to attempt the moves uphill, and hence increasingly penalized by their Boltzman factor.

A simple interpretation of the success of the $c^* = 0$ strategy is as follows. If we assume that the next move would have negligibly different Δ from the current one, relative to the sampling error σ and hence in terms of move acceptance, then c^* should correspond to the threshold at which the acceptance probability of the next sample would correspond to that of the first sample of a new move. The latter begins with $c_0 = 0$ so this match occurs at $c_n = c^* = 0$. In Appendix B we note there is some depth to this idea: the further one proceeds down the decision chain without accepting or rejecting the move, the more the posterior distribution for Δ narrows around $\Delta = 0$.

Given that this optimal value is independent of the particular value of Δ, why is it not then globally optimal for an arbitrary distribution of Δ? The answer is that it is less effective than the strategy of the preceding section at persevering with moves which are relatively favorable within the spread of the prior distribution. It is also informative to ask what happens if we (almost) never reject, which amounts to trying to maximize the acceptance per move rather than per sample. At large n it helps to view the cumulating cost changes c_n as a random walk (in one dimension) with drift: it has random steps of standard deviation σ but non-zero mean Δ. For the zero drift case

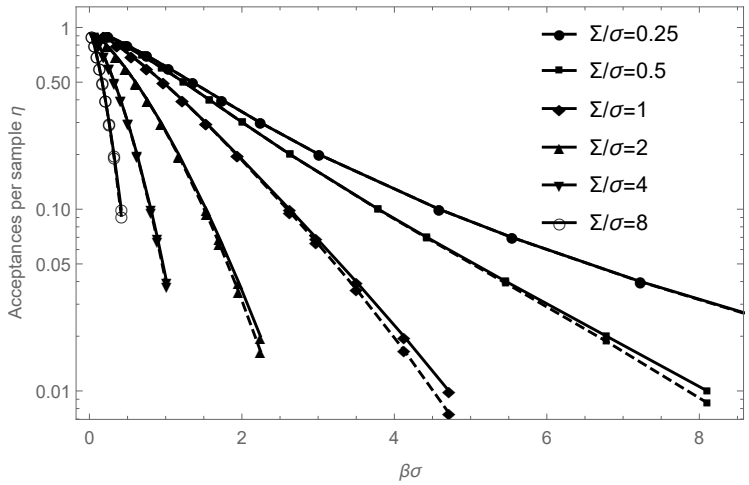


Figure 1 **Performance in terms of move acceptance per sample, for the optimal rejection strategy (solid lines) compared to the much simpler $c^* = 0$ rejection strategy (dashed lines). All of these results use the optimal acceptance rule derived in Section 5.2, and it can be seen that the reduction in performance due to using the simpler rejection rule (dashed lines) is small across the wide range of parameters plotted. The parameters Σ and σ are the respective standard deviations of the dethermalized prior distribution of underlying performance difference Δ and of the errors superposed on these when sampling, whilst β is the reciprocal temperature of the thermal annealing.**

$\Delta = 0$ such a random walk is recurrent to (the region of) its origin, which will always eventually lead to move acceptance by crossing $c + \beta\sigma^2/2 = 0$. With non-zero drift the process is no longer recurrent at large n and it will be dominated by the linear increase in the drift term, $c_n \simeq n\Delta$ at large n. For positive drift $\Delta > 0$ this means the process can escape towards $c = +\infty$ never to return, and hence with non zero probability the move is never accepted. If we incorporate move rejection at some large positive c^* then these never accepted instances eventually become rejected, and because we have $P_A(\Delta) = 1$ for $\Delta \leq 0$ the resulting move acceptance probability must then be that of Metropolis (6). Whilst $c^* \to \infty$ achieves Metropolis, it takes far too many samples before rejecting a move and thus, in terms of samples per move, is not practical. At the other extreme setting $c^* = -\beta\sigma^2/2$ leads to no continuation beyond the first step and our decision chain reduces to the single step decision of Ceperly and Dewing, which rejects very many moves which again leads to a very large number of samples per accepted move.

5.4. Summary and Procedural Algorithm

We end this section by summarizing what we have and have not achieved. Ideally we would like to have a decision chain constrained to obey detailed balance only at the level of each proposed move, and which is optimal in its trade-off of move acceptance vs. sample cost. We have imposed the stronger constraint of detailed balance at each sample step, which reduces to Equation (18) for which we find a universally optimal acceptance rule (19). This acceptance rule is remarkably

universal in the sense that it achieves the absolute highest possible chance of move acceptance at each step, and this is independent of the choice for the move rejection decisions. We proposed an optimal rejection threshold under the assumption that the Δ of the following move is the same as for the current move, and will demonstrate in the following section that this works well in practice.

Procedurally, the algorithm thus works as follows, stopping at the first *decision:*

1. *Accept* with probability $A = \min\left(1, e^{-2(c_n+\beta\sigma^2/2)(c_{n-1}+\beta\sigma^2/2)/\sigma^2}\right)$; else
2. If $c_n > c^*$, *Reject*; else
3. Continue and take another sample.

In the following, this algorithm with $c^* = 0$ is denoted as Optimized Stochastic Annealing (OSA). Note that, with respect to the categories introduced in Section 3, OSA modifies the acceptance criterion but obeys the detailed balance equation, and uses fully sequential sampling to adjust the number of samples per evaluation until it is confident enough to make a decision, while it does not require to keep track of all samples of all visited solutions as some of the other approaches.

6. Performance Comparison

In this section, we compare the OSA method derived above with two other methods that promise to follow a set annealing schedule and to obey the detailed balance equation, i.e., SANE and CD described in Section 4.

The first set of experiments will investigate the algorithm's response to a particular quality difference Δ and thus be independent of an optimization problem, neighborhood definition or annealing schedule. The second set will then compare optimization performance on a standard traveling salesperson problem (TSP) where the quality (tour length) of a solution is disturbed by adding a Gaussian noise whenever a solution is evaluated. In all cases, when we report on solution quality, we refer to the true (undisturbed) quality of the current solution.

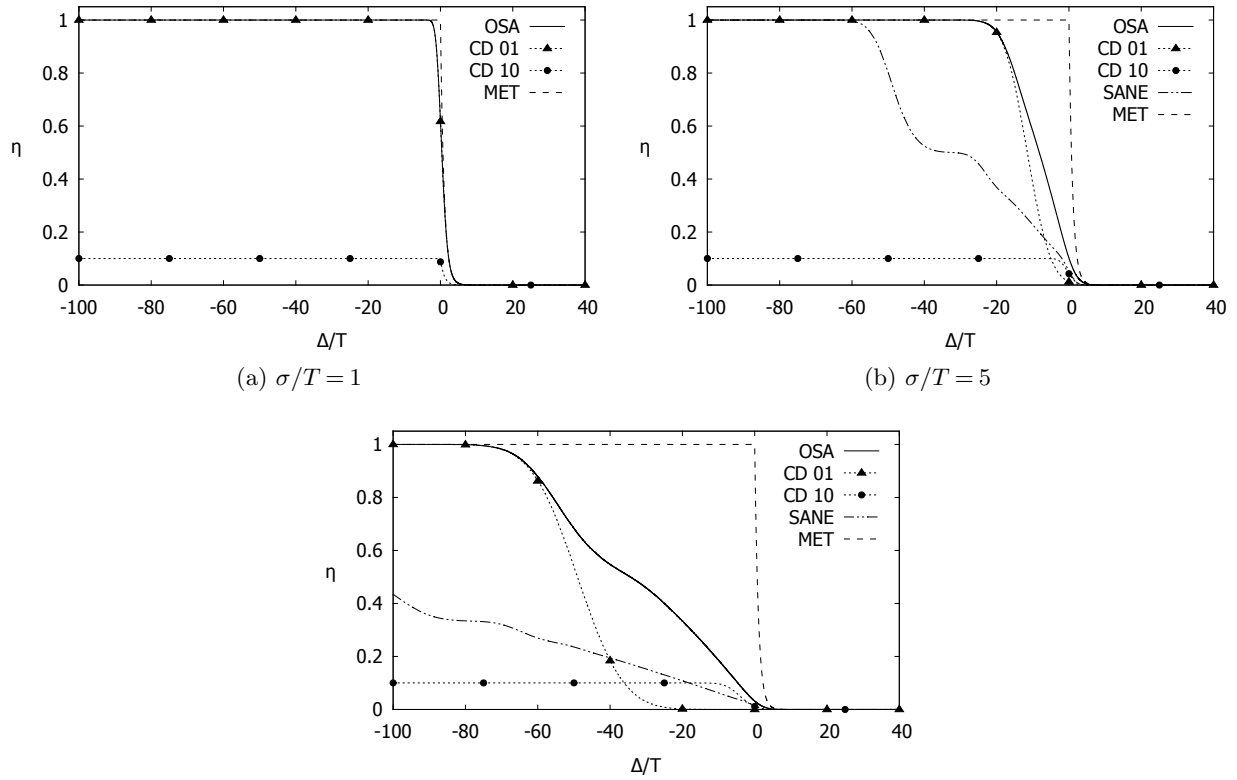
6.1. Comparison of Acceptance Curves

As explained above, the speed at which the algorithm converges to equilibrium depends on the rate of accepting moves. The key performance indicator is thus the expected computational effort required before accepting a move. Assuming that the computational effort is primarily depending on the number of samples taken, we would therefore prefer algorithms which have a higher acceptance probability per sample, which we defined as η in Equation (12).

The acceptance probability depends on Δ, σ and T. However, all of Δ, σ and T are given in arbitrary (but common) units of cost (or energy) so the resulting algorithm performance can only depend on two independent dimensionless combinations which we choose as Δ/T and σ/T.

Figure 2 shows the effective probability of accepting a move with a given normalized quality difference divided by the expected number of samples required to make a decision for different

ratios σ/T. The curves were computed empirically, by subjecting the algorithm 100,000 times to a true quality difference of Δ, and recording how often it would accept the quality difference in the end and how many samples it took to make a decision, for the given variance of the noise.



(a) $\sigma/T = 1$ (b) $\sigma/T = 5$

(c) $\sigma/T = 10$

Figure 2 **Overall algorithm performance in terms of effective acceptance probability per sample, for different σ/T. For $\sigma/T = 1$, SANE works identical to CD01. Metropolis corresponds to $\sigma/T = 0$ (the deterministic case) and is only shown for reference.**

As can be seen, for low σ/T, most methods very closely follow the Metropolis criterion, although they cannot replicate the sharp bend that Metropolis has at $\Delta/T = 0$. However, because CD10 always needs 10 samples to make a decision, its acceptance probability per sample never exceeds 0.1, which clearly shows that taking 10 samples is a waste of computational effort for low σ/T.

As σ/T is increased and the problems become more challenging, as expected, the acceptance probability per sample for all methods drops. CD01 remains relatively steep, but moves to the left, which means that even for improving moves, unless the improvement is rather large, the acceptance probability is virtually zero. CD10 is better for smaller improvements, but as explained before, can never reach an acceptance probability per sample of more than 0.1, even if the improvement is very large.

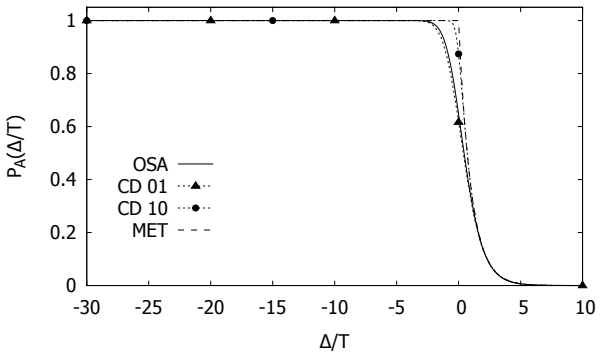
The curve for SANE basically becomes more stretched towards the left, so it has a higher acceptance probability than CD01 for small improvements, and a lower acceptance probability for medium improvements, but still approaches acceptance probability per sample of 1 for very large improvements. The "steps" in the curve of SANE comes from the fact that the number of samples collected is integer. If the new solution is much better, one sample is always sufficient to detect this and accept the solution. For smaller Δ/T, a single sample is normally not sufficient to be confident about accepting the new solution, and thus a second sample is required, which takes the acceptance probability per sample down to 0.5. Similarly, the next steps can be seen at 1/3, 1/4, etc.

OSA has the *highest* acceptance probability per sample in *all* three scenarios, for *all* Δ/T. It is particularly noteworthy that its curve is always above CD01 and CD10, which shows that, while following similar ideas as CD01 and CD10, it is able to adjust the number of samples appropriately to the situation at hand. Similar to SANE and for the same reasons, the curve for OSA also shows some steps, although to a lesser extent.

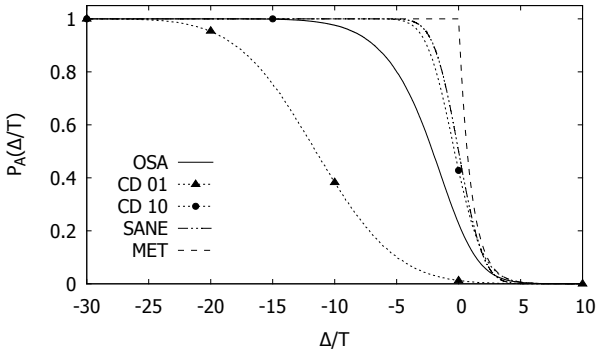
Figures 3 and 4 shed more light on how the algorithms achieve this performance by separately considering the acceptance probability per attempted move and the number of samples before a decision can be made.

Figure 3 shows the effective probability of accepting a move depending on the actual normalized quality difference Δ/T for different σ/T. For small ratio σ/T (Figure 3 a), all curves look similar, they more or less follow the Metropolis criterion. CD10 can replicate Metropolis slightly better as it effectively has to deal with a reduced noise. The situation changes as σ/T is increased (Figures 3 b,c). SANE now has the highest acceptance probability, and it is noteworthy that it is almost unaffected by changes to σ/T. The other methods increasingly move away from the Metropolis criterion without noise. In other words, also new solutions that are substantially better than the current solution now risk being rejected. The effect is most dramatic for CD01 which, at $\sigma/T = 10$ shows acceptance probabilities of less than 0.02 for $\Delta/T > -30$. This explains why, as we will see later, CD basically stops accepting new solutions and stalls as σ is increased or the temperature is lowered.

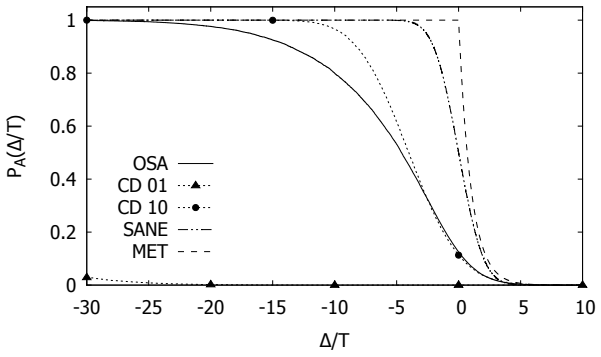
The information about the number of samples an algorithm requires to make a decision is shown in Figure 4. As CD always uses a fixed number samples, its lines are flat, whereas the sequential procedures (SANE and OSA) take more samples when the solutions have a similar quality, and less samples when there is a large quality difference and the better solution can be identified easily (all curves approach no. of samples $= 1$ as $|\Delta/T| \to \infty$). Intuitively this makes sense, as sampling effort is channelled to situations where the deterministic acceptance functions have a high gradient, and a wrong estimation of Δ would have a large impact on the acceptance decision, whereas for large



(a) $\sigma/T = 1$

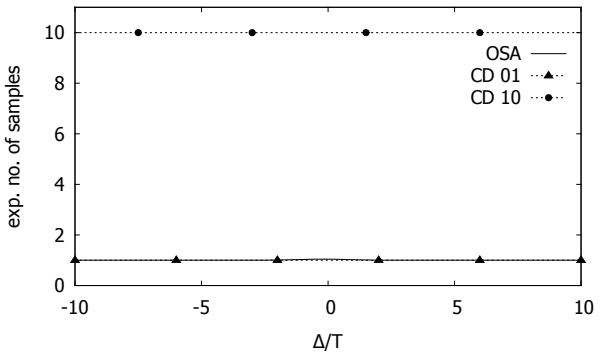


(b) $\sigma/T = 5$

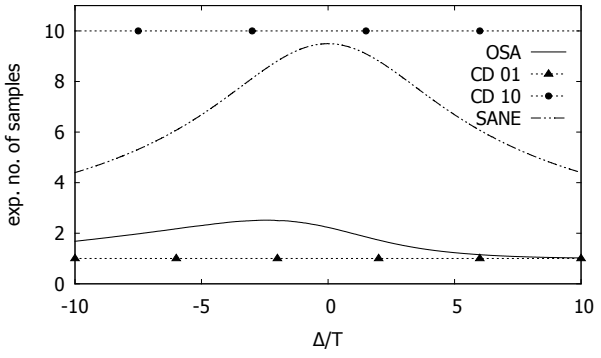


(c) $\sigma/T = 10$

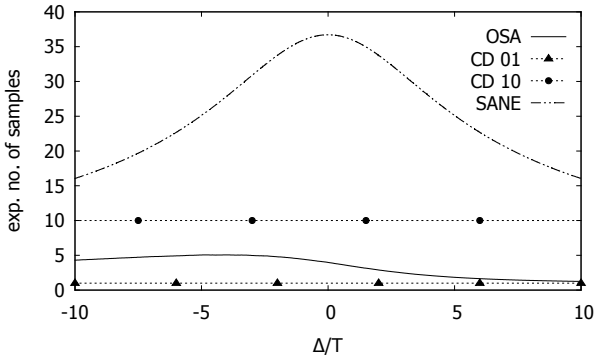
Figure 3 **Effective acceptance probability curves $P_A(\Delta/T)$ of different methods for different ratios σ/T. The Metropolis criterion (MET) is depicted without noise for reference. Note the different methods require different levels of computational effort to arrive at each decision (see Figure 2 on the right).**



(a) $\sigma/T = 1$



(b) $\sigma/T = 5$



(c) $\sigma/T = 10$

Figure 4 **Expected number of samples before an accept or reject decision is made, for different σ/T**

$|\Delta|$ the acceptance curves are flat and therefore accuracy is less important. The figures also show that SANE and OSA are able to automatically increase the number of samples taken for larger

ratios of σ/T, i.e., where decisions are more difficult or accuracy of decisions is more important. In general, SANE seems to require more samples to make a decision than OSA. The asymmetric decision policy of OSA is also clearly visible in these plots.

6.2. Convergence to Equilibrium

Now let us validate empirically whether the methods really work as promised and obey the detailed balance equation. To this end, we run each algorithm with a fixed temperature $T = 15$ on a TSP problem with Gaussian noise of zero mean and $\sigma^2 = 200$ added to the objective function. The particular TSP instance we use is `eil51` taken from TSPLIB (2002). On the algorithm side, we use a standard permutation representation with swap neighborhood (exchange of two randomly selected cities in the permutation).

If the algorithms indeed obey the detailed balance equation, they should all converge to the same average solution quality as Metropolis and Glauber in the deterministic case. As can be seen in Figure 5, showing the solution quality over the number of samples used, this is indeed the case for OSA and the CD variants as would expected from theory. SANE, on the other hand, converges to a substantially lower average cost which is clearly well below equilibrium at the temperature set. It is also clearly visible that all CD runs, while converging to the correct solution quality level, are significantly slower than OSA, i.e., they are clearly less efficient, at least on this temperature level. Among the three CD variants, for this temperature level, CD01 is the fastest, followed by CD10 and then CD100 which requires about 50,000 samples to reach equilibrium (outside of the figure). This further strengthens the conclusions drawn from Figure 2.

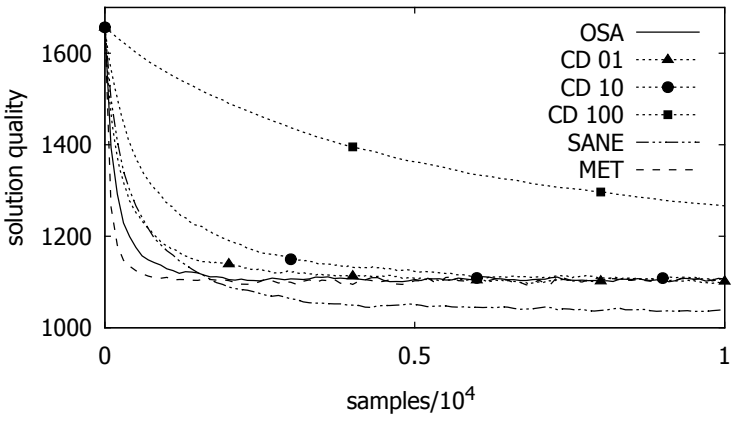


Figure 5 Solution quality over samples for the different algorithms. Results have been averaged over 500 runs.

6.3. Comparison over Optimization

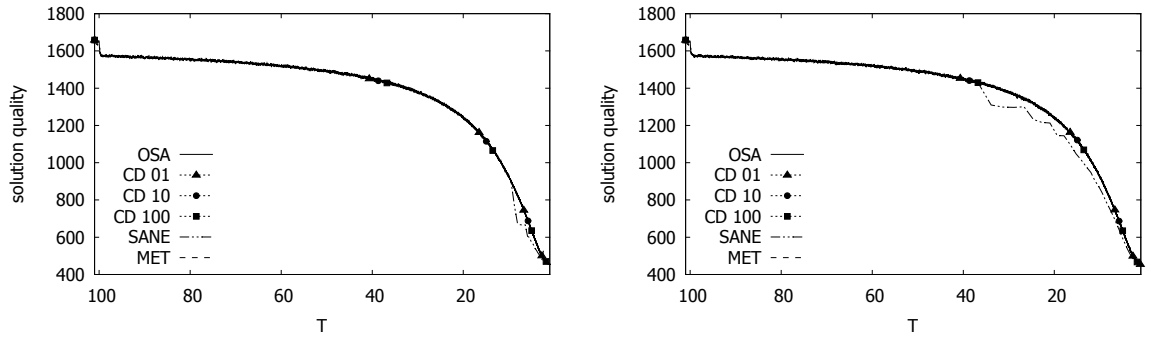
Now let us consider optimization performance on the *eil51* TSP problem. When we compare the different algorithms on an actual optimization task, we need to decide on an annealing schedule. However, since all methods at least aim (and all but SANE actually achieve) to obey the detailed balance equation, if we give each method the same number of accepted moves on each temperature stage, they all have the same chance to approach equilibrium at each temperature level, and they should all roughly reach the same solution quality. They will still differ, however, in their efficiency of actually making acceptance moves, i.e., in the number of samples they require at each temperature stage, allowing us to demonstrate the *efficiency* gain of our derived fully sequential method, while removing the influence of the annealing schedule as much as possible.

Since we furthermore believe that lowering the temperature smoothly is better than lowering it occasionally in big steps, we have designed a geometric annealing schedule with an initial temperature of $T_0 = 100$ which is lowered by multiplying it with $\alpha = 0.9999$ after every *accepted* move. This setting of α has been determined by tuning the annealing scheme for the Metropolis algorithm on the deterministic version of the problem.

All results below are averaged over 1000 runs with different random seeds. We use common random numbers to sharpen contrasts between the different procedures. Since for all tested methods the computational overhead can be considered small compared to the time required for a single evaluation, running time is equated with the number of evaluations performed.

Figure 6 shows the quality of the current solution at each temperature level during the run (note that because we reduce the temperature after every accepted move, we have exactly one current solution for each temperature level and each run). The figure verifies our assumption that all algorithms achieve more or less the same solution quality for each temperature level. Also, by comparing Figures 6 (a) and (b), we see that the methods achieve the same solution qualities irrespective of the noise level, i.e., they effectively eliminate the noise. The biggest deviation can be observed for SANE, as expected, since we know from the experiments in Section 6.2 that it is not exactly obeying the detailed balance equation. However, we still regard this close enough to justify using the same annealing schedule for all approaches.

The solution quality obtained depending on runtime (number of samples taken) can be observed in Figure 7. At larger variance of $\sigma^2 = 3200$, the differences between the algorithms are more pronounced. Overall, it is clearly visible that OSA performs best, converging to the best solution quality most quickly. SANE is second best, confirming the good performance reported by Branke et al. (2008). While the advantage of OSA over SANE is small (but significant) in the case of $\sigma^2 = 200$, OSA is substantially faster in the case of higher variance. Comparing the three versions of CD, CD01 converges fastest, but then stalls. As explained above, this is due to the fact that it



(a) $\sigma^2 = 200$ (b) $\sigma^2 = 3200$

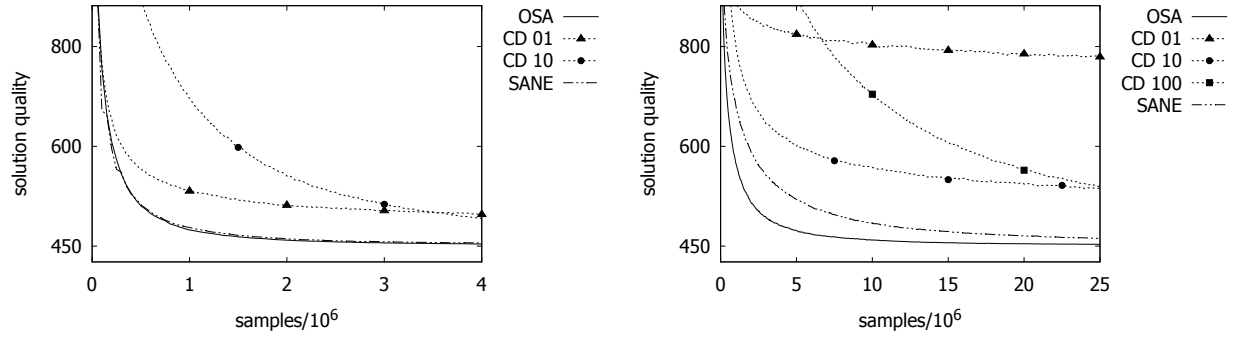
Figure 6 Average solution quality encountered at temperature T for small and large noise level.

Table 1 Comparison of optimization results for different computational budgets and noise levels. All differences are statistically significant according to a paired Wilcoxon rank sum test with significance level 0.05.

	$\sigma^2 = 200$			$\sigma^2 = 3200$		
samples/10^6	1	2.5	4	5	15	25
OSA	471.3	455.2	452.9	470.5	454.4	452.3
SANE	474.8	456.9	454.0	514.7	469.5	460.2
CD01	527.7	502.0	493.2	829.2	791.5	776.7
CD10	688.8	525.8	487.8	600.1	545.0	532.3
CD100	1,470.3	1,222.0	1013.1	925.2	606.7	534.5

stops accepting even improving moves once the temperature is low. CD10 and CD100 can delay this effect by essentially lowering the noise level, and thus can converge to better final solutions but at the expense of significantly slower convergence.

Table 1 summarizes the achieved solution quality of the different methods after different computational budgets. According to the Wilcoxon rank sum test with significance level 0.05, OSA significantly outperforms SANE and CD in all cases.

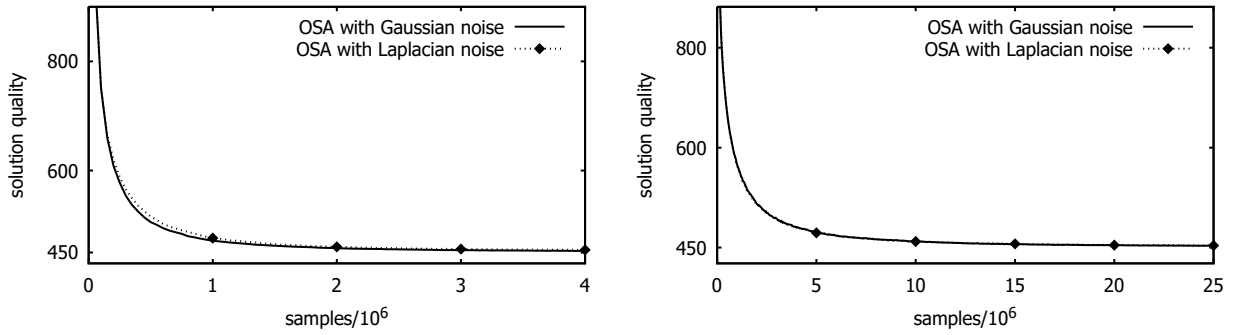


(a) $\sigma^2 = 200$ (b) $\sigma^2 = 3200$

Figure 7 Solution quality achieved depending on the number of samples used, for different noise levels.

6.4. Robustness to non-Gaussian Noise

Simulation output is often normally distributed due to the central limit theorem, because it is the result of many small influences or explicitly an average value (such as the average flow time of jobs in a production process). If the output is not normally distributed, one option may be to batch samples, and make acceptance or rejection decisions not after every sample, but after every batch of samples. But in any case, OSA should have a high tolerance with respect to non-normal distributions, in particular in the case of high noise, where many samples are usually taken before a solution is accepted or rejected. In this section, we examine the effect of non-normal noise on the performance by running experiments with a Laplace (double exponential) distribution. Figure 8 compares the optimization performance of OSA under a Gaussian noise with that under the Laplace noise having identical variance. As can be seen, there is a tiny difference in the low variance case, especially in the early phase of optimization where temperature is still high and the number of samples taken low. In the high variance case, as expected, the performance under the two noise distributions is indistinguishable. This confirms the robustness of OSA to non-Gaussian noise distributions.



(a) $\sigma^2 = 200$ (b) $\sigma^2 = 3200$

Figure 8 **Solution quality achieved depending on the number of samples used, for Gaussian and Laplace noise distributions.**

7. Conclusions

We have proposed the Optimized Stochastic Annealing (OSA) algorithm, an efficient variant of Simulated Annealing that obeys the detailed balance equation despite noise in the evaluation function, assuming the noise is Gaussian with known variance. The algorithm is fully sequential in the sense that it decides after every sample either to accept or reject the new solution in accordance with detailed balance, or to continue with another sample. We show that under conditions of detailed balance at each step, acceptance and rejection decisions can be made independently, and derive an acceptance rule that is both optimal in the sense that it maximizes the acceptance probability per

sample, which is equivalent to maximizing the convergence rate to thermodynamical equilibrium, and also universal in that it is independent of the distribution of moves. For the rejection criterion, we find the optimal rule and show how it depends on expectations of possible alternative moves. We also propose using a simpler rejection threshold that is optimal under the assumption that the performance difference of the next move is identical to the performance difference of the current move, and almost as efficient as the optimal rejection rule in a wide range of scenarios. We have empirically compared OSA with two other state-of-the-art algorithms, namely one by Ceperley and Dewing (1999) which also obeys the detailed balance equation, and a recent algorithm by Branke et al. (2008) that has been shown to outperform several other SA variants. First, we confirmed that OSA has indeed a higher acceptance probability per sample than the other two algorithms. Second, we have empirically compared the algorithms on a TSP problem with noisy objective function, and shown that OSA converges to better solutions more quickly than the other two algorithms. OSA's efficiency make it an excellent candidate for simulation optimization, where the number of samples that can be taken is limited because of the time required to run simulations.

There are various ways to extend this work, such as relaxing the assumption of known Gaussian noise, and making better use of samples gathered in previous iterations to fine tune the rejection threshold.

Appendix A: Derivation of Universal Optimal Acceptance Rule

To optimize each of the acceptance and rejection rules we will have recourse to calculating backwards up the decision chain, from a terminating step at $n = N$ beyond which the move is rejected without further consideration. The loss of efficiency η can be made as small as we like by choosing N large enough, as follows. We first note that the probability $u(n)$ of a move remaining undecided after n decision steps must decrease faster than $1/n$, because otherwise the mean number of steps per decision which can be expressed as $n_s = \sum_n u(n)$ would diverge and the efficiency η would be zero. It follows that if we choose to terminate the decision chain after the N'th step we can make the loss in overall move acceptance P_A as small as we like, whilst the corresponding denominator of n_s in (12) does not increase.

To implement detailed balance at each step we first need to note that the Gaussian form we assume for the sample error distribution f can be factorized in its dependence on Δ as

$$f(c_{i+1} - c_i - \Delta) = e^{c_{i+1}\Delta/\sigma^2} f(c_{i+1} - c_i) e^{-c_i\Delta/\sigma^2} e^{-\Delta^2/(2\sigma^2)} \tag{25}$$

and this in turn leads to the probability density of chain histories continuing (at least) as far as the n'th observation being given from Equation (16) by:

$$\begin{aligned} P_n(C_n|\Delta) &= e^{c_n\Delta/\sigma^2 - n\,\Delta^2/(2\sigma^2)} \prod_{i=1}^{n-1} \left(f(c_{i+1} - c_i) S(C_i)\right)\, f(c_1) \\ &= e^{c_n\Delta/\sigma^2 - n\,\Delta^2/(2\sigma^2)} P_n(C_n|0). \end{aligned} \tag{26}$$

What is crucial about the above expression is that all of the dependence of P_n on Δ and most particularly the sign of Δ is in the simple front factor. Substituting the expression (17) for $P_a(\Delta, n)$ at each step in (18) and using the explicit dependence (26) of $P_n(C_n|\Delta)$ on Δ then expresses detailed balance as the requirement that the following be an even function in terms of Δ,

$$\int A(C_n) e^{\Delta(c_n+\beta\sigma^2/2)/\sigma^2 - n\,\Delta^2/(2\sigma^2)} P_n(C_n|0) H(\pm(c_{n-1}+\beta\sigma^2/2)) dc_1 dc_2 \ldots dc_n = \text{even}(\Delta), \tag{27}$$

where $H(x) = (1+\text{sgn}(x))/2$ is the Heaviside function selecting for the respective sign cases of condition (18).

Forward from this point notation and symmetry of argument are more clear substituting c_n in terms of the translated variables

$$z_n = c_n + \beta\sigma^2/2. \tag{28}$$

The LHS of the above can then be written as $e^{-n\,\Delta^2/(2\sigma^2)}\tilde{F}(\Delta)$ in terms of the two sided Laplace transform $\tilde{F}(\Delta) = \int_{-\infty}^{\infty} e^{z_n \Delta} F_{\pm}(z_n) dz_n$, which is the transform of $F_{\pm}(z_n) = \int A(C_n) P_n(C_n|0) H(\pm z_{n-1}) dz_1 .. dz_{n-1}$. The two sided Laplace transform has the property that even functions (here of z_n) transform to even functions (here of Δ) and likewise odd to odd as in the equivalent Fourier Transform result, e.g. Bracewell (1965). It follows that for expression (27) to be even in Δ it is equivalent that $F_{\pm}(z_n)$ be an even function of z_n. To examine this let us denote $A(C_n) = A(z_n, C_{n-1})$ and substitute $P_n(C_n|0)$ in terms of $P_{n-1}(C_{n-1}|0)$ to convert the detailed balance conditions (18) to

$$F_{\pm}(z_n) = \int A(z_n, C_{n-1}) e^{z_n z_{n-1}/\sigma^2} e^{-(z_n^2+z_{n-1}^2)/(2\sigma^2)} S(C_{n-1}) P_{n-1}(C_{n-1}|0) H(\pm z_{n-1}) dz_1 .. dz_{n-1} = \text{even}(z_n). \tag{29}$$

Note that all dependence on the sign of z_n comes from the first two factors of the integrand, $A(z_n, C_{n-1}) e^{z_n z_{n-1}/\sigma^2}$, so one possible way to satisfy Equation (29) is to impose

$$A(z_n, C_{n-1}) e^{z_n z_{n-1}/\sigma^2} = e^{-z_n z_{n-1}/\sigma^2} A(-z_n, C_{n-1}), \tag{30}$$

which we will see later does prove optimal.

We now seek to choose the acceptance rules $A(z_n, C_{n-1})$ which will maximize the move acceptance per sample (12), constrained by (29). We start by maximizing the acceptance at the end of the chain ($n = N$) because there is no possible offsetting later gain beyond. When $z_{n-1} > 0$ we can choose $A(z_n, C_{n-1}) = 1$ for all $z_n < 0$, which clearly makes $F_+(z_n < 0)$ as large as possible, and then ensure $F_+(z_n) = F_+(-z_n)$ by using (30) to obtain that for $z_n > 0$, $A(z_n, C_{n-1}) = e^{-2 z_n z_{n-1}/\sigma^2}$. Note that we have thereby maximized $F_+(z_n)$ for all z_n under the constraint that this function be even. For $z_{n-1} < 0$ a parallel argument applies based on accepting with probability 1 all cases where $z_n > 0$, and leads to the same acceptance expressions except that the applicability ranges of z_n are reversed. All of this can be collected together as

$$A(z_n, C_{n-1}) = A(z_n, z_{n-1}) = \min\left(1, e^{-2 z_n z_{n-1}/\sigma^2}\right) \tag{31}$$

which has further key properties which we will subsequently exploit in inferring the applicability of the same form for successively lower $n < N$. Foremost it gives detailed balance at the n'th step within the set of fixed

prior history C_{n-1}, corresponding the *integrands* in Equation (29) being even in z_n, with the important consequence that the acceptance probability of a given history C_{n-1} depends on that history alone: different histories do not interact.

We now consider the optimization of the acceptance rule for step n' under the assumption that we have already chosen the rule (31) for all steps $n > n'$. The non-interaction of histories in later steps means that it cannot pay to forgo any possible acceptance probability at this step also, as any gain in deferred acceptance cannot exceed that forgone and moreover it would come at higher sampling cost. Hence we deduce that the same acceptance rule (31) is optimal at step n' and by induction we find that it applies at all steps. Another key property of rule (31) is that the decision chain can only transmit (to later decisions) z_n having same sign as z_{n-1}, because other cases are certain to be accepted and exit the decision chain. Thus given the initial condition $z_0 = c_0 + \frac{1}{2}\beta\sigma^2 = \frac{1}{2}\beta\sigma^2 > 0$ we end up with no decisions which inherit $z_{n-1} < 0$ at any level.

Appendix B: Evaluation of Optimal Rejection Strategy

As in Appendix A, we will work in terms of translated cumulative cost variables $z_n = c_n + \beta\sigma^2/2$. The expectations in our recurrence relations (21) and (22) then have to be taken with respect to z_{n+1} conditional on the value of z_n, including how the latter influences the distribution of Δ for a decision chain which has gone this far. The resulting distribution we denote by $P(z_{n+1}|z_n)$ and we find it as follows. We start by assuming that we have obtained an estimate of the distribution $P_0(\Delta)$ of Δ over all attempted moves, which for any new move is the distribution prior to any samples being drawn. Now we further require the posterior distribution of Δ for any particular decision chain after it has taken n (known) samples. By Bayes' theorem this is given by $P(\Delta|C_n) = P(C_n|\Delta)P_0(\Delta)/P(C_n) \propto e^{-\frac{1}{2}\sum_{i=1}^{n}(\delta_i-\Delta)^2/\sigma^2}P_0(\Delta)$. Expanding the square in the exponent, noting $\sum_{i=1}^{n} y_i = z_n - \frac{\beta}{2}\sigma^2$ and normalising the distribution then leads to

$$P(\Delta|z_n) = P_0(\Delta)e^{-\frac{n}{2}\Delta^2/\sigma^2 + \Delta(z_n/\sigma^2 - \beta/2)}/Q_n(z_n) \tag{32}$$

where the normalising denominator is given by

$$Q_n(z_n) = \int P_0(\Delta)e^{-\frac{n}{2}\Delta^2/\sigma^2 + \Delta(z_n/\sigma^2 - \beta/2)}d\Delta. \tag{33}$$

We can now write $P(z_{n+1}|z_n) = \int P(z_{n+1}|z_n, \Delta)P(\Delta|z_n)d\Delta$ where $P(z_{n+1}|z_n, \Delta) = f(z_{n+1} - z_n - \Delta)$ as in (11) and $P(\Delta|z_n)$ is given by Equation (32), and after integrating over Δ this simplifies down to

$$P(z_{n+1}|z_n) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}(z_{n+1}-z_n)^2/\sigma^2}\frac{Q_{n+1}(z_{n+1})}{Q_n(z_n)}. \tag{34}$$

It is worth noting that for a decision chain to reach large n the magnitude of Δ must be small for the chain not to have drifted into an accept or reject decision. This is reflected in the form of the posterior $P(\Delta|z_n)$ given by Equation (32) which will tend to $e^{-\frac{n}{2}\Delta^2/\sigma^2}$ as $n \to \infty$ and the $Q_n(z_n)$ defined in Equation (33) becomes correspondingly insensitive to z_n and with $Q_{n+1}/Q_n \to 1$. Thus all decision chains asymptotically enter a common regime where the posterior Δ is effectively zero.

We can now render more explicit the averages in the recursion relations (21) and (22) using the distribution (34) and deduce how to initialize our recurrence relations. In particular the expected step acceptance is given by

$$\mathbb{E}(A(z_{n+1}, z_n)) = \int_{-\infty}^{\infty} \frac{dz_{n+1}}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}(z_{n+1}-z_n)^2/\sigma^2}\frac{Q_{n+1}(z_{n+1})}{Q_n(z_n)}\min\left(1, e^{-2z_{n+1}z_n/\sigma^2}\right)$$

which in the limit of $n \to \infty$ so that $Q_{n+1}(z_{n+1})/Q_n(z_n) \to 1$ reduces to $\mathbb{E}(A(z_{n+1}, z_n))_{n\to\infty} = \text{erfc}\left(\frac{z_n}{\sigma\sqrt{2}}\right)$, and likewise

$$\mathbb{E}(S(z_{n+1}, z_n)q_{n+2}(z_{n+1})) = \int_0^{z^*_{n+1}} \frac{dz_{n+1}}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(z_{n+1}-z_n)^2/\sigma^2} \frac{Q_{n+1}(z_{n+1})}{Q_n(z_n)} \left(1 - e^{-2z_{n+1}z_n/\sigma^2}\right) q_{n+2}(z_{n+1}) \quad (35)$$

with $q_{n+2} \equiv a_{n+2}$, s_{n+2} respectively. The recurrence relations can now in principle be initialized by first positing some very large $n = N$ beyond which all is rejected, equivalent to imposing $z^*_N = 0$: for large enough N this will entail negligible loss in algorithm performance as discussed at teh beginning of Appendix A, and we can also use the large n limit to eliminate Q_N/Q_{N-1}. Then $a_N(z_{N-1}) = \text{erfc}\left(\frac{z_{N-1}}{\sigma\sqrt{2}}\right)$, $s_N(z_{N-1}) = 1$ and there are no contributions from further continuation so we can infer that z^*_{N-1} obeys $\text{erfc}\left(\frac{z^*_{N-1}}{\sigma\sqrt{2}}\right) = \eta$. Given all of these we can then find all the corresponding expressions for $n = N-2$ from the full recurrence relations and iterate all the way back to $n = 1$ and self-consistency for η.

Our recurrence relations assume no particular form for the prior distribution of Δ for attempted moves except that it be known. However in thermal equilibrium we must have $P_0(-\Delta) = e^{-\beta\Delta}P_0(\Delta)$, because the states from which the moves are attempted differ in energy by Δ and hence correspondingly in equilibrium probability. It is therefore natural to consider the "equilibrium priors" as

$$P_0(\Delta) = e^{+\beta\Delta/2}P_{00}(|\Delta|), \quad (36)$$

where the dethermalized prior $P_{00}(\Delta)$ will be symmetric. Most conveniently it turns out that the way the prior enters the recurrence relations is in terms of the dethermalized prior through the definition (33) of $Q_n(z_n)$, which then simplifies to

$$Q_n(z_n) = \int P_{00}(\Delta)e^{-\frac{n}{2}\Delta^2/\sigma^2 + z_n\Delta/\sigma^2} d\Delta, \quad (37)$$

and now has no explicit dependence on β. For computations of the decision chain parameters, we can then take η determining the rejection thresholds (20) as truly an input parameter, and then the closing condition (23) as identifying the corresponding value of β which gives this. This is exploited in the numerical results shown in Section 5.3.

References

Ahmed MA, Alkhamis TM (2002) Simulation-based optimization using simulated annealing with ranking and selection. *Computers and Operations Research* 29(4):387–402.

Alkhamis T, Ahmed M, Tuan V (1999) Simulated annealing for discrete optimization with estimation. *European Journal of Operational Research* 116(3):530–544.

Alkhamis TM, Ahmed MA (2004) Simulation-based optimization using simulated annealing with confidence interval. Ingalls RG, et al., eds., *Winter Simulation Conference*, 514–519.

Alrefaei M, Andradóttir S (1999) A simulated annealing algorithm with constant temperature for discrete stochastic optimization. *Management Science* 45(5):748–764.

Alrefaei MH, Diabat AH (2009) A simulated annealing technique for multi-objective optimization. *Applied Mathematics and Computation* 215(8):3029–3035.

Avello E, Baesler FF, Moraga RJ (2004) A meta-heuristic based on simulated annealing for solving multiple-objective problems in simulation optimization. *Winter Simulation Conference*, 508–513.

Ball R, Fink T, Bowler N (2003a) Characterization of the probabilistic traveling salesman problem. *Physical Review E* 68(3):036703.

Ball R, Fink T, Bowler N (2003b) Stochastic annealing. *Physical Review Letters* 91(3):030201.

Bowler NE (2001) *The role of noise in optimisation and diffusion limited aggregation.* Ph.D. thesis, University of Warwick.

Bracewell RN (1965) *The Fourier Transform and Its Applications* (New York: McGraw Hill).

Branke J, Meisel S, Schmidt C (2008) Simulated annealing in the presence of noise. *Journal of Heuristics* 14:627–654.

Bulgak A, Sanders J (1988) Integrating a modified simulated annealing algorithm with the simulation of a manufacturing system to optimize buffer sizes in automatic assembly systems. *Winter Simulation Conference*, 684–690 (IEEE).

Ceperley D, Dewing M (1999) The penalty method for random walks with uncertain energies. *Journal of Chemical Physics* 110(20):9812–9820.

Cohn H, Fielding M (1999) Simulated annealing: Searching for an optimal temperature schedule. *SIAM Journal of Optimization* i:779–802.

Fink T (1998) *Inverse Protein Folding, Hierarchical Optimisation and Tie Knots.* Ph.D. thesis, University of Cambridge.

Fox BL, Heine GW (1995) Probabilistic search with overrides. *Annals of Applied Probability* 5(4):1087–1094.

Gelfand S, Mitter S (1989) Simulated annealing with noisy or imprecise energy measurements. *Journal of Optimization Theory and Applications* 62(1):49–62.

Glauber R (1963) Time-dependent statistics of the ising model. *Journal of Mathematical Physics* 4(2):294–307.

Gutjahr W, Pflug G (1996) Simulated annealing for noisy cost functions. *Journal of Global Optimization* 8(1):1–13.

Gutjahr WJ (2005) Two metaheuristics for multiobjective stochastic combinatorial optimization. *Stochastic Algorithms: Foundations and Applications*, 116–125, LNCS (Springer).

Hastings W (1970) Monte carlo sampling methods using markov chains and their applications. *Biometrika* 57(1):97–109.

Johnson DS, Aragon CR, McGeoch LA, Schevon C (1989) Optimization by simulated annealing: An experimental evaluation; part I, graph partitioning. *Operations Research* 37:865–892.

Johnson DS, Aragon CR, McGeoch LA, Schevon C (1991) Optimization by simulated annealing: An experimental evaluation; part II, graph coloring and number partitioning. *Operations Research* 39:378–406.

Kennedy A, Kuti J (1985) Noise without noise: A new monte carlo method. *Physical Review Letters* 54(23):2473–2476.

Kirkpatrick S, Gelatt C, Vecchi M (1983) Optimization by simulated annealing. *Science* 220(4598):671–680.

Kushner H (1987) Asymptotic global behavior for stochastic approximation and diffusions with slowly decreasing noise effects: Global minimization via monte carlo. *SIAM Journal on Applied Mathematics* 47(1):169–185.

Mattila V, Virtanen K, Hämäläinen RP (2012) A simulated annealing algorithm for noisy multiobjective optimization. *Journal of Multi-Criteria Decision Analysis* 20(5-6):255–276.

Metropolis N, Rosenbluth A, Rosenbluth M, Teller A, Teller E (1953) Equation of state calculations by fast computing machines. *Journal of Chemical Physics* 21(6):1087–1092.

Nourani Y, Andresen B (1998) A comparison of simulated annealing cooling strategies. *Journal of Physics A: Mathematical and General* 31:8373–8385.

Painton L, Diwekar U (1995) Stochastic annealing for synthesis under uncertainty. *European Journal of Operational Research* 83(3):489–502.

Peskun P (1973) Optimum monte-carlo sampling using markov chains. *Biometrika* 60(3):607–612.

Prudius AA, Andradottir S (2005) Two simulated annealing algorithms for noisy objective functions. Kuhl ME, et al., eds., *Winter Simulation Conference*, 797–802 (IEEE).

Prudius AA, Andradottir S (2012) Averaging frameworks for simulation optimization with applications to simulated annealing. *Naval Research Logistics* 59(6):411–429.

Rees S, Ball R (1987) Criteria for an optimum simulated annealing schedule for problems of the travelling salesman type. *Journal of Physics A: Mathematical and General* 20(5):1239–1249.

Romeo F, Sangiovanni-Vincentelli AL (1991) A theoretical framework for simulated annealing. *Algorithmica* 6:302–345.

TSPLIB (2002) http://www.iwr.uni-heidelberg.de/groups/comopt/software/TSPLIB95/index.html.

van Laarhoven PJM, Aarts EHL (1987) *Simulated Annealing: Theory and Applications* (Reidel Publishing).

Wang L, Zhang L (2006) Stochastic optimization using simulated annealing with hypothesis test. *Applied Mathematics and Computation* 174(2):1329–1342.